

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

"EXPRESS MAIL" LABEL NUMBER: EV 118 993 375 US

DATE OF DEPOSIT: January 9, 2004

I HEREBY CERTIFY THAT THIS PAPER OR FEE IS BEING DEPOSITED WITH THE UNITED STATES POSTAL SERVICE "EXPRESS MAIL POST OFFICE TO ADDRESSEE" SERVICE UNDER 37 CFR § 1.10 ON THE DATE INDICATED ABOVE AND IS ADDRESSED TO: MAIL STOP PATENT APPLICATION, COMMISSIONER FOR PATENTS, P.O. BOX 1450, ALEXANDRIA, VA 22313-1450.

SAUL SOLANO

(TYPED OR PRINTED NAME OF PERSON MAILING PAPER)

(SIGNATURE OF PERSON MAILING PAPER OR FEE)

UTILITY
APPLICATION

for

UNITED STATES LETTERS PATENT

on

COMPOSITIONS AND METHODS FOR DETERMINING CANINE GENDER

by

Bonita J. M. Ferrie
Eric V. Johnston
Sue K. DeNise

Docket No.: MMI1130-1
Drawings: Six Sheets

Lisa Haile, JD, Ph.D.
GRAY CARY WARE & FREIDENRICH LLP
4365 Executive Drive, Suite 1100
San Diego, California 92121-2133

COMPOSITIONS AND METHODS FOR DETERMINING CANINE GENDER

CROSS REFERENCE TO RELATED APPLICATION

[0001] This application claims the benefit of priority under 35 U.S.C. § 119(e) of U.S. Serial No. 60/439,188, filed January, 10, 2003, the entire content of which is incorporated herein by reference.

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

[0002] The invention relates generally to genetic analysis and more specifically to genetic analysis related to canine gender.

BACKGROUND INFORMATION

[0003] DNA-based techniques have become accepted methods for the determination of gender. The use of DNA for gender determination has been particularly effective in situations where sex is difficult to determine accurately by visual or manual examination, such as in certain avian species, and in circumstances where visual inspection is not possible, and gender must be determined from a biological sample alone.

[0004] For example, Griffiths, *et al.* (1996), U.S. Pat. No. 5,508,165 (Halverson *et al.* '165), U.S. Pat. No. 5,679,514 (Baker), and U.S. Pat. No. 5,707,809 (Halverson *et al.* '809) disclose methods for determining the sex of many avian species by DNA techniques. These methods amplify fragments specific to avian sex chromosomes. In some other species, chromosomal regions corresponding to the amelogenin gene have been used to develop DNA-based sexing tests. Amelogenin is involved in the production of tooth enamel and in many mammals there exists a copy of the gene on the X chromosome (AMELX) and another on the Y chromosome (AMELY). Differences in these homologs have allowed the development of gender tests in which different sized fragments are simultaneously amplified from X and Y chromosomes using a single set of primers. These tests have been developed for human (Sullivan *et al.* 1993), equine (Hasegawa *et al.* 2000) and bovine DNA (Chen *et al.* 1999). Although in some cases these primers are not 100% specific to the species being sex-typed and will amplify fragments from other species (Buel *et al.* 1995),

they are unique enough to be only useful as gender markers for the species for which they are developed. The markers developed for human gender testing, for example, do not distinguish male and female canines.

[0005] The gender determination of an individual dog using DNA-based techniques would be an important quality control step in applications that utilize canine genotyping. Most canine registries request that information about an individual dog, including gender, be provided by the owner upon registration. In some cases DNA samples, in the form of cheek swabs, are also provided for identity determination or parentage verification. Since the individual dog is not available for visual inspection, the determination of gender from DNA isolated from the submitted sample would verify accuracy of the submitted gender information by the registrant, and improves the integrity of data that becomes part of the registry database.

SUMMARY OF THE INVENTION

[0006] The present invention provides a method to determine gender of a canine subject, that includes contacting a nucleic acid sample from the canine subject with at least one probe or primer specific for canine amelogenin, and using the binding of the at least one probe or primer to detect a difference between the canine amelogenin gene on the Y chromosome and the canine amelogenin gene on the X chromosome, thereby determining gender of the canine subject. In certain aspects, binding of the at least one probe or primer is detected. For example, gender of the canine subject can be determined by contacting the nucleic acid sample with at least one probe or primer that specifically binds SEQ ID NO:22 and/or SEQ ID NO:23.

[0007] In certain aspects, gender of the canine subject is determined by contacting the nucleic acid sample with at least one probe or primer that specifically binds SEQ ID NO:10 (Canine X amelogenin sequences 1-143 (Table 3)) and/or SEQ ID NO:11. (Canine Y amelogenin sequences 1-146 (Table 3)). For example, a first primer of a primer pair specifically binds to SEQ ID NO:6 (Canine X amelogenin sequences 1-76 (Table 3)) and/or SEQ ID NO:7 (Canine Y amelogenin sequences 1-70 (Table 3)), and a second primer of the

primer pair specifically binds to SEQ ID NO:8 (Canine X amelogenin sequences 77-143 (Table 3)) and/or SEQ ID NO:9 (Canine Y amelogenin sequences 71-146 (Table 3)). In certain aspects, the first primer is SEQ ID NO:3 or SEQ ID NO:4, and the second primer is SEQ ID NO:5.

[0008] In certain aspects and embodiments disclosed herein, in addition to detecting binding of at least one probe or primer to a canine amelogenin gene, methods of the present invention include detecting binding of at least one probe or primer to a canine microsatellite locus, to provide genotype and gender information.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] Figure 1 illustrates determination of gender in canine samples by the analysis of nucleic acid products derived from the use of canine amelogenin specific primers. The plots, generated using an ABI Prism, shows X and Y specific amplification products for various canine samples, generated using BFK90059 and BFK90065. The X axis is scan number and the Y axis is relative intensity of signal.

[0010] Figure 2 illustrates multiplex PCR results for canine samples, wherein the multiplex PCR assay includes a gender marker and 13 microsatellite markers (i.e. STRs). The plots, generated using an ABI Prism, denote the various markers, including an X specific amplification product generated using BFK90059 and BFK90065. X and Y axis are as disclosed for Figure 1. The plots were generated from a sample of a female canine subject using different colored labels, as indicated.

[0011] Figure 3 illustrates multiplex PCR results for canine samples, wherein the multiplex PCR assay includes a gender marker and 13 microsatellite markers (i.e. STRs). The plots, generated using an ABI Prism, denote the various markers, including X and Y-specific amplification products (i.e. markers) generated using BFK90059 and BFK90065. X and Y axis are as disclosed for Figure 1. The plots were generated from a male subject using different colored labels, as indicated.

[0012] Figure 4 illustrates multiplex PCR results for canine samples, wherein the multiplex PCR assay includes a gender marker and 4 microsatellite markers (i.e. STRs). The plots, generated using an ABI Prism, denote the various markers, including X and Y-specific amplification products (i.e. markers) generated using BFK90059 and BFK90065. X and Y axis are as disclosed for Figure 1. The plots were generated from a sample from a male canine subject, using different colored labels, as indicated.

[0013] Figure 5 provides a comparison of consensus sequences of canine X (SEQ ID NO:22) and Y (SEQ ID NO:23) partial amelogenin sequences. The shaded sequences and gaps indicate differences between the canine AMELX and AMELY sequences.

DETAILED DESCRIPTION OF THE INVENTION

[0014] The present invention is based on the discovery of differences in the canine amelogenin gene homologs on the X (AMELX) and Y (AMELY) chromosomes. Furthermore, the present invention is based on the design and identification of PCR primers that specifically amplify regions on both canine sex chromosomes and accurately distinguish them, thereby providing amplification products that are canine gender markers. The PCR primers can be used in multiplexing genotyping assays. The primers that were identified to amplify and distinguish canine sex chromosomes were designed so that they do not amplify human DNA, a possible source of contamination in canine genotyping analysis. Therefore, these primers are useful as quality control standards in canine genotyping reactions, to assure that amplification products are generated from canine template polynucleotides and not polynucleotides from a technician.

[0015] Canine gender determination is an important component of canine parentage and identity testing and provides quality assurance when used in conjunction with DNA-based parentage and identity testing. The results of canine gender determination are used by breed registries in creating pedigrees for specific animals. These data are added to breed registry databases and thereby become a component of an animal's official pedigree. Further, results can be used in forensic cases involving canines, whereby results could be used to

"include or exclude" specific animals in a manner similar to human forensic genotyping practices. The gender markers can also serve as stand alone markers, or as part of a set of standard markers for comparison of genotypes.

[0016] Accordingly, provided herein is a method to determine gender of a canine subject, that includes contacting a nucleic acid sample from the canine subject with at least one probe or primer specific for canine amelogenin, and detecting binding of the at least one probe or primer, thereby determining gender of the canine subject. Binding of the at least one probe or primer distinguishes the canine amelogenin gene on the X chromosome from the amelogenin gene on the Y chromosome. The presence of a copy of the amelogenin gene on the Y chromosome indicates that the subject is male, whereas the absence of a copy of the amelogenin gene on the Y chromosome indicates that the subject is female, thereby providing a determination of the gender of the canine subject.

[0017] In certain aspects, gender of the canine subject is determined by contacting the nucleic acid sample with at least one probe or primer that specifically binds SEQ ID NO:22, a region of the canine amelogenin gene on the X chromosome, and/or SEQ ID NO:23, a corresponding region of the canine amelogenin gene on the Y chromosome, and which are effective at allowing differences between the sequence of SEQ ID NO:22 and/or SEQ ID NO:23 to be identified.

[0018] The probe or primers specific for the canine amelogenin gene are designed so that they can be used to detect differences disclosed herein between the canine amelogenin gene on the X chromosome and the canine amelogenin gene on the Y chromosome. These differences are apparent, for example, from a comparison of SEQ ID NO:20, the nucleotide sequence of a region of the canine amelogenin gene on the X chromosome, and SEQ ID NO:21, the nucleotide sequence of the corresponding region of the canine amelogenin gene on the Y chromosome. It will be understood based on the differences in these sequences, that probes and primers can be designed that only recognize the X or Y version, or that can be used to generate a product that is detectably distinguishable between the X and Y version of the gene. Figure 5 provides a comparison of larger partial sequences of the X (SEQ ID NO:22) and Y (SEQ ID NO:23) copies of the amelogenin gene. It will be

understood that based on the present disclosure, specific assays can be devised that exploit the differences in the X and Y chromosome versions of the canine amelogenin gene, for example the partial sequences of the X and Y chromosome versions of the amelogenin gene provided in SEQ ID NOS:22 and 23.

[0019] Furthermore, it will also be understood that the sequence of the entire canine (e.g. dog) amelogenin gene can be identified and characterized using standard methods based on the present disclosure, combined with known sequences of the amelogenin gene in other species, for example the human amelogenin gene sequence data (Genbank accession numbers AY0040206 (genomic) and M86932 (mRNA)(*Homo sapiens* amelogenin X (AMELX)); and Genbank accession numbers AC013412 (genomic (Bac clone RP11-507A3)) and M86933(mRNA)(*Homo sapiens* amelogenin Y (AMELY)), and dog genomic data (The Dog Genome: Survey Sequencing and Comparative Analysis, Kirkness et al. (2003) *Science*, Vol 301 1898-1903, incorporated by reference herein in its entirety (Genbank accession numbers AACN010000001-AACN011089636 and CE000001-CE853796)). Additional regions of the canine amelogenin gene are likely to be identified from the complete canine amelogenin gene sequence that include nucleotide sequence differences between the copy of the gene on the X chromosome and that on the Y chromosome. These additional regions can be utilized to design additional probes and primers that can be used in the methods disclosed herein, in place of, or in combination with, the probes and primers disclosed herein. Furthermore, based on the present disclosure, it will be understood that it is likely that other canine genes that are found on both the X and Y chromosomes will include X or Y chromosome-specific differences that can be exploited in methods to determine canine gender, for example the Y chromosome SRY gene and the X homologue SOX3, and in the ZFY and ZFX genes (Griffiths, R. (2000) Sex identification using DNA markers; In: Molecular methods in ecology (Ed. by Baker AJ) Blackwell Science, London).

[0020] A method of the present invention to determine gender of a canine subject, can include contacting a nucleic acid sample from the canine subject with at least one probe or primer specific for a canine amelogenin gene, and detecting binding of the at least one

probe or primer to a target region of the canine amelogenin gene and/or detecting an extension product of the canine amelogenin gene generated from the at least one primer. A difference in the binding of the at least one probe or primer, or a difference in the extension product of the at least one probe or primer, distinguishes the copy of the canine amelogenin gene on the X chromosome from the copy of the canine amelogenin gene on the Y chromosome. The presence of a copy of the amelogenin gene on the Y chromosome indicates that the subject is male, whereas the absence of a copy of the amelogenin gene on the Y chromosome indicates that the subject is female, thereby providing a determination of the gender of the canine subject. The target region is close enough to a nucleotide(s) that is different between amelogenin gene on the X chromosome and the amelogenin gene on the Y chromosome, such that binding of the probe or primer can be used to detect the difference between the amelogenin gene on the X chromosome from the amelogenin gene on the Y chromosome. For example, the target region can be within 1000, 500, 250, 150, 100, 50, 25, 20, 15, 10, 5, 4, 3, 2, or 1 nucleotide of a nucleotide position that is different between the amelogenin gene on the X chromosome and the amelogenin gene on the Y chromosome.

[0021] A method to determine gender of a canine subject of the present invention, can include contacting a nucleic acid sample from the canine subject with at least one probe or primer specific for canine amelogenin, and detecting a canine amelogenin gene on a Y chromosome based on binding of the probe or primer. The presence of a copy of the amelogenin gene on the Y chromosome indicates that the subject is male, whereas the absence of a copy of the amelogenin gene on the Y chromosome indicates that the subject is female, thereby providing a determination of the gender of the canine subject. In certain aspects, the methods involve detecting both the canine amelogenin gene X chromosome copy and the canine amelogenin gene Y chromosome copy to determine gender of the canine subject.

[0022] In another related embodiment, the present invention provides a method to determine gender of a canine subject, wherein a nucleic acid sample from the canine subject is contacted with at least one probe or primer specific for canine amelogenin, and the binding of the at least one probe or primer is used to detect a difference in the canine

amelogenin copy on the Y chromosome and the canine amelogenin copy on the X chromosome, thereby determining gender of the canine subject.

[0023] In general, methods of the present invention determine whether a subject is male or female (i.e. determine the gender of a canine subject) by attempting to detect the presence of a copy of the amelogenin gene on the Y chromosome, which indicates that the subject is male. Conversely, the absence of a copy of the amelogenin gene on the Y chromosome indicates that the subject is female. The method can also include detecting a copy of the amelogenin gene on the X chromosome. The detection of a copy of the amelogenin gene on the X chromosome, but not the amelogenin gene on the Y chromosome, is determinative of a female subject. The detection of a copy of the amelogenin gene on the X chromosome and a copy of the amelogenin gene on the Y chromosome, is determinative of a male subject. The detection of the presence of the amelogenin gene on the X chromosome, in addition to providing gender information, provides further assurance that an assay was successful. In certain aspects, as indicated herein, the method detects only the presence of the amelogenin gene on the Y chromosome.

[0024] In certain aspects of methods disclosed herein, binding of at least one probe or primer is detected and used to determine gender of the canine subject by determining the size of an amplification product(s) generated using the primer or primer pair, or by determining the size of a restriction fragment(s) generated from a sample and identified using the probe. For example, probes and/or primers can be used to synthesize or detect a nucleic acid product that is different between the canine amelogenin gene copy on the X chromosome and the canine amelogenin gene copy on the Y chromosome. The nucleic acid can be, for example, a primer extension product (e.g. a sequencing reaction product), an amplification product, or a restriction enzyme cleavage product. As illustrated in Figure 1, methods disclosed herein can detect the presence or absence of the Y chromosome by detecting the presence or absence an amplification product that is detectably distinguishable depending on whether it is generated from the Y chromosome or the X chromosome.

[0025] It will be understood based on the present disclosure that other methods can be used to exploit the canine sequence information provided herein to distinguish the X from

the Y chromosome copy of the amelogenin gene. These methods include, but are not limited to, polynucleotide sequencing, primer extension, differential hybridization, oligo ligation assay (OLA), mass spectroscopy or a SNP detection method, of which many are known in the art.

[0026] In certain aspects, gender of the canine subject is determined by contacting the nucleic acid sample with at least one probe or primer that specifically binds SEQ ID NO:10 (Canine X amelogenin sequences 1-143 (Table 3)) and/or SEQ ID NO:11. (Canine Y amelogenin sequences 1-146 (Table 3)). For example, at least one probe or primer can specifically bind to SEQ ID NO:6 (Canine X amelogenin sequences 1-76 (Table 3)) and/or SEQ ID NO:7 (Canine Y amelogenin sequences 1-70 (Table 3)), and at least one other probe or primer can specifically bind to SEQ ID NO:8 (Canine X amelogenin sequences 77-143 (Table 3)) and/or SEQ ID NO:9 (Canine Y amelogenin sequences 71-146 (Table 3)). Such methods take advantage of the X or Y chromosome-specific insertions and/or deletions between SEQ ID NO:10 and SEQ ID NO:11.

[0027] In certain aspects, a primer pair is employed in the methods for determining canine gender. In fact, primer pairs used in such methods represent another embodiment of the invention. In methods disclosed herein that employ a primer pair, typically the first primer and the second primer bind opposite strands of a target polynucleotide and prime polynucleotide synthesis from the target polynucleotide in opposite directions. Typically, the primer pairs are used in an amplification reaction, as are known in the art and disclosed in more detail herein. In certain embodiments, the present invention provides an isolated primer pair that primes polynucleotide synthesis of the canine amelogenin gene across a region that includes an X or Y chromosome-specific insertion or deletion.

[0028] Accordingly, in these aspects gender of the canine subject is determined by contacting the nucleic acid sample with a primer pair. Target sequences for binding of primers of the primer pair can be sequences that are at least 80%, 85%, 90%, 95%, 99%, or 100% identical between the canine X and Y amelogenin genes and bind (i.e. hybridize) to target regions of the canine amelogenin gene that flank insertion/deletion areas. These primer pairs are designed to produce amplification fragments of different sizes for the X and

Y chromosomes. In certain aspects, one primer of the primer pair is labeled. Detectable labeling of a polynucleotide or oligonucleotide such as a primer, is well known in the art, as discussed in more detail herein. In certain aspects, for example, a probe or primer is fluorescently labeled.

[0029] In certain examples, a first primer of the primer pair specifically binds to SEQ ID NO:6 and/or SEQ ID NO:7 and a second primer of the primer pair specifically binds to SEQ ID NO:8 and/or SEQ ID NO:9. A specific illustration of this is provided in the Example provided herein. One set of primers (BFK90052 (SEQ ID NO:5) and BFK90055 (SEQ ID NO:3) produces a 143 base band in male and female canine samples and an extra band of 146 bases in male samples. Another set of primers, SEQ ID NO:5 and SEQ ID NO:4 produces a 139 base band in male and female canine samples and an extra band of 142 bases in male samples. It will be understood that other regions of the amelogenin gene that flank the region of the gene identified in Table 3 (SEQ ID NO:10 and SEQ ID NO:11) can be used as binding regions for a primer pair to amplify across this region, which includes insertions and deletion differences between the canine X and canine Y chromosomes.

[0030] Furthermore, typically the primer pairs used in the methods of the present invention generate different sized amplification products in canines than in humans, or generate lower levels or no amplification products with human samples. For example, as illustrated in the Example, when human DNA was tested using primers BFK90065 (SEQ ID NO:4) and BFK90059 (SEQ ID NO:5), a faint 143 base band was detected in female and male but no 146 base band was detected in any human sample, thus eliminating any concern of human DNA contamination. These results were expected since BFK90065 is in a highly conserved region near the initial amelogenin ATG start codon, whereas BFK90059 lies within an intron and has 19/22 bases identical to human amelogenin-X but only 15/22 bases identical to human amelogenin-Y with the 3' end of the primer being most different (Table 1).

[0031] In certain aspects, the first primer includes at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22 contiguous nucleotides of SEQ ID NO:3, and the second

primer includes at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22 contiguous nucleotides of SEQ ID NO:5. In another aspect, the first primer includes at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or 18 contiguous nucleotides of SEQ ID NO:4, and the second primer includes at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22 contiguous nucleotides of SEQ ID NO:5. In another aspect, the first primer is at least 90%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO:3 or SEQ ID NO:4, and the second primer is at least 90%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO:5.

[0032] The exact composition of the primer or probe sequences of the present methods are not critical to the invention, but they must hybridize to sequences of the canine amelogenin gene under moderate or high stringency conditions (i.e. specifically bind), or for probes or primers to microsatellite sequences, they must hybridize under moderate or high stringency conditions to U and/or U' sequences that flank repeats of the microsatellites, as disclosed herein. Therefore, a probe or primer that is specific for a canine amelogenin gene, binds to the canine amelogenin gene under moderate or high stringency conditions. Furthermore, a probe or primer that is specific for a canine amelogenin gene is identified such that they are useful for performing the methods disclosed herein. For example, when a primer pair that is specific for a canine amelogenin gene is used to amplify a region of the canine amelogenin gene, an amplification product is generated that is a unique size from any other detectable amplification products generated using the primer pair to amplify canine (e.g. dog) genomic DNA. To maximize the resolution of size differences between products generated from the amelogenin gene on the X and Y chromosome or from a microsatellite locus, it is preferable to choose a primer sequence that is close to an insertion or deletion sequence for the amelogenin gene, or close to the repeat sequence for the microsatellites. Thus, primers usually bind within at least about 100 nucleotides, 50 nucleotides, or 25 nucleotides of an insertion, deletion, or repeat.

[0033] Algorithms for the selection of primer sequences are generally known and are available in commercial software packages. Primers can be designed to be used in primer extension reactions. Primer pairs of the present invention can be designed to hybridize to complementary strands of chromosomal DNA, and prime towards an insertion, deletion, or

repeat sequences, so that these sequences are amplified to produce a product of a detectably different length between individuals and/or for the X and Y chromosome. Primers will usually be at least about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or 18 nucleotides in length, and usually not more than about 50, 45, 40, 35, 30, or 25 nucleotides in length.

[0034] It will be recognized that some amount of non-specific hybridization is unavoidable, but is acceptable provided that hybridization to a target nucleotide sequence is sufficiently selective such that it can be distinguished over the non-specific cross-hybridization, for example, at least about 2-fold more selective, generally at least about 3-fold more selective, usually at least about 5-fold more selective, and particularly at least about 10-fold more selective, as determined, for example, by an amount of labeled oligonucleotide that binds to target nucleic acid molecule as compared to a nucleic acid molecule other than the target molecule, particularly a substantially similar (i.e., homologous) nucleic acid molecule other than the target nucleic acid molecule. Conditions that allow for selective hybridization can be determined empirically, or can be estimated based, for example, on the relative GC:AT content of the hybridizing oligonucleotide and the sequence to which it is to hybridize, the length of the hybridizing oligonucleotide, and the number, if any, of mismatches between the oligonucleotide and sequence to which it is to hybridize (see, for example, Sambrook et al., "Molecular Cloning: A laboratory manual (Cold Spring Harbor Laboratory Press 1989)).

[0035] An example of progressively higher stringency conditions is as follows: 2 x SSC/0.1% SDS at about room temperature (hybridization conditions); 0.2 x SSC/0.1% SDS at about room temperature (low stringency conditions); 0.2 x SSC/0.1% SDS at about 42°C (moderate stringency conditions); and 0.1 x SSC at about 68°C (high stringency conditions). Washing can be carried out using only one of these conditions, e.g., high stringency conditions, or each of the conditions can be used, e.g., for 10-15 minutes each, in the order listed above, repeating any or all of the steps listed. However, as mentioned above, optimal conditions will vary, depending on the particular hybridization reaction involved, and can be determined empirically.

[0036] The term "polynucleotide" is used broadly herein to mean a sequence of deoxyribonucleotides or ribonucleotides that are linked together by a phosphodiester bond. For convenience, the term "oligonucleotide" is used herein to refer to a polynucleotide that is used as a primer or a probe.

[0037] An oligonucleotide or polynucleotide can be RNA or can be DNA, which can be genomic DNA, a gene or a portion thereof, a cDNA, a synthetic polydeoxyribonucleic acid sequence, or the like, and can be single stranded or double stranded, as well as a DNA/RNA hybrid. In various embodiments, a polynucleotide, including an oligonucleotide (e.g., a probe or a primer) can contain nucleoside or nucleotide analogs, or a backbone bond other than a phosphodiester bond. In general, the nucleotides comprising a polynucleotide or oligonucleotide are naturally occurring deoxyribonucleotides, such as adenine, cytosine, guanine or thymine linked to 2'-deoxyribose, or ribonucleotides such as adenine, cytosine, guanine or uracil linked to ribose. However, a polynucleotide or oligonucleotide also can contain nucleotide analogs, including non-naturally occurring synthetic nucleotides or modified naturally occurring nucleotides. Such nucleotide analogs are well known in the art and commercially available, as are polynucleotides containing such nucleotide analogs (Lin et al., *Nucleic Acids Research* (1994) 22:5220-5234 Jellinek et al., *Biochemistry* (1995) 34:11363-11372; Pagratis et al., *Nature Biotechnol.* (1997) 15:68-73, each of which is incorporated herein by reference). Oligonucleotides, such as primers and probes, can also be comprised of peptide nucleic acids (PNA) (Nielsen PE and Egholm M. (1999) *Curr. Issues Mol. Biol.* 1:89-104).

[0038] The covalent bond linking the nucleotides of an oligonucleotide or polynucleotide generally is a phosphodiester bond. However, the covalent bond also can be any of numerous other bonds, including a thiodiester bond, a phosphorothioate bond, a peptide-like bond or any other bond known to those in the art as useful for linking nucleotides to produce synthetic polynucleotides (see, for example, Tam et al., *Nucl. Acids Res.* (1994) 22:977-986, Ecker and Crooke, *BioTechnology* (1995) 13:351360, each of which is incorporated herein by reference). The incorporation of non-naturally occurring nucleotide analogs or bonds linking the nucleotides or analogs can be particularly useful where the

oligonucleotide or polynucleotide is to be exposed to an environment that can contain a nucleolytic activity, including, for example, a tissue culture medium or upon administration to a living subject, since the modified polynucleotides can be less susceptible to degradation.

[0039] A polynucleotide or oligonucleotide that includes naturally occurring nucleotides and phosphodiester bonds can be chemically synthesized or can be produced using recombinant DNA methods, using an appropriate polynucleotide as a template. In comparison, a polynucleotide or oligonucleotide comprising nucleotide analogs or covalent bonds other than phosphodiester bonds generally are chemically synthesized, although an enzyme such as T7 polymerase can incorporate certain types of nucleotide analogs into a polynucleotide and, therefore, can be used to produce such a polynucleotide recombinantly from an appropriate template (Jellinek et al., *supra*, 1995). Thus, the term polynucleotide as used herein includes naturally occurring nucleic acid molecules, which can be isolated from a cell, as well as synthetic molecules, which can be prepared, for example, by methods of chemical synthesis or by enzymatic methods such as by the polymerase chain reaction (PCR).

[0040] A method of the present invention can be performed using a specific binding pair member to a canine gender maker. As used herein, the term "specific binding pair member" refers to a molecule that specifically binds or selectively hybridizes to another member of a specific binding pair. Specific binding pair member include, for example, probes, primers, polynucleotides, antibodies, etc. For example, a specific binding pair member includes a primer or a probe that selectively hybridizes to a target polynucleotide that includes an amelogenin gene, or that specifically hybridizes to an amplification product generated using the target polynucleotide as a template.

[0041] In certain aspects, a probe is employed in the methods for determining canine gender of the present invention. Accordingly, in these aspects gender of the canine subject is determined by contacting the nucleic acid sample with a probe that specifically binds a target sequence. The probe is typically labeled, for example fluorescently labeled. Target sequences for binding of primers of the primer pair can include, for example, sequences

within the canine amelogenin genes that include insertion/deletion areas, with the aim of obtaining differential binding of the probe to the X and Y chromosomes. For example, a probe can be designed to bind to a target region that includes nucleotides 55 to 59 of the canine amelogenin gene present on the X chromosome (i.e. nucleotides 55 to 59 of SEQ ID NO:14 and SEQ ID NO:22), such that the probes do not bind to a corresponding target region of the amelogenin gene on the Y chromosome.

[0042] As used herein, the term "at least one", when used in reference to a primer, probe, microsatellite, or the like, means 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, etc. Reference to "at least a second" primer, probe, microsatellite, or the like, means two or more, i.e., 2, 3, 4, 5, 6, 7, 8, 9, 10, etc., primers probes, microsatellites, or the like.

[0043] As used herein, the term "infer" or "inferring", when used in reference to gender, means drawing a conclusion about gender using a process of analyzing individually or in combination, amplification products or probe-bound polynucleotides generated from a nucleic acid sample of the subject, using a method disclosed herein, and comparing the amplification products or probe-bound polynucleotides to known gender makers. Typically, for methods of the invention, a qualitative analysis, such as identifying whether 1 or 2 amplification products are produced, is sufficient to infer the gender of the subject.

[0044] A canine subject analyzed using the methods disclosed herein is a member of the family Canidae, which includes the species dogs, wolves, coyotes, foxes and jackals. The degree of homology between these species is high, and the subject methods are expected to be effective with any canine species. Furthermore, as illustrated herein, the methods have accurately determined gender of a subject in a wide range of dog breeds, including many common breeds of dogs. The nucleic acid from the canine subject is typically a genomic DNA sample that is isolated from the subject or subjects that are to be tested. DNA can be isolated from any nucleated cellular source such as blood, hair shafts, saliva, mucous, biopsy, or feces. Amplification can be performed on the DNA from a single cell, although it is convenient to use at least about 10^5 cells.

[0045] Typically, primers of the present invention are used to amplify a region of genomic DNA that includes an insertion or deletion that is present in the X or Y

chromosome, but not both the X and Y chromosome. The amplification reaction can be, as non-limiting examples, the polymerase chain reaction or the ligase chain reaction. Suitable reaction conditions for PCR are described, for example, in Saiki, et al. (1985) *Science* 239:487, and Sambrook, et al. *Molecular Cloning: A Laboratory Manual*, CSH Press 1989, pp.14.2-14.33. Useful thermostable polymerases known in the art include those isolated, for example, from *Thermus aquaticus*, *Thermococcus litoralis*, *Pyrococcus furiosus*, and *Thermus thermophilus*, as is known. A discussion of the ligase chain reaction (LCR) may be found in International patent application WO 9302215.

[0046] Conveniently, a probe or primer of the present invention can include a detectable label. Suitable labels include fluorochromes, e.g. fluorescein isothiocyanate (FITC), rhodamine, Texas Red, phycoerythrin, allophycocyanin, 6-carboxyfluorescein (6-FAM), 2',7'-dimethoxy-4',5'-dichloro-6-carboxyfluorescein (JOE), 6-carboxy-X-rhodamine (ROX), 6-carboxy-2',4',7',4,7-hexachlorofluorescein (HEX), 5-carboxyfluorescein (5-FAM) or N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA), radioactive labels, e.g. ^{32}P , ^{35}S , or ^3H . The label can be a two stage system, where the amplified DNA is conjugated to biotin or haptens, for example, that have a high affinity binding partner, for example avidin or specific antibodies, where the binding partner is conjugated to a detectable label. The label can be conjugated to one or both of the primers. Alternatively, the pool of nucleotides used in the amplification can be labeled, so as to incorporate the label into the amplification product.

[0047] After amplification, the products can be size fractionated. Fractionation can be performed, for example, by gel electrophoresis, such as denaturing acrylamide or agarose gels. A convenient system uses denaturing polyacrylamide gels in combination with an automated DNA sequencer, see Hunkapillar et al. (1991) *Science* 254:59-74. The automated sequencer is particularly useful with multiplex amplification. Capillary electrophoresis can also be used for fractionation. A review of capillary electrophoresis can be found in Landers, et al. (1993) *BioTechniques* 14:98-111.

[0048] In one aspect, reaction and PCR conditions for a method to determine gender provided herein, include the following: 0.29 pmoles of each of BFK90059 and BFK90065

primers, 1.5X Taq Gold Buffer, 2.5 mM MgCl₂, 0.175 mM each dNTP and 0.75 units of AmpliTaq Gold Polymerase in an 5 µl reaction containing a highly variable amount of canine DNA (approximately 0.5-100 ng). PCR: 95°C for 12 minutes, 15 cycles of 30 seconds at 94°C, 30 seconds at 64°C decreasing 0.2°C/cycle to 61°C, and 1 minute at 72°C, followed by 20 cycles of 30 seconds at 94°C, 30 seconds at 61°C, and 1 minute at 72°C, followed by a final extension of 72°C for 45 minutes in a Perkin Elmer 9600 thermocycler.

[0049] In another embodiment, the present invention provides an isolated oligonucleotide that can be used as a primer or probe in the methods disclosed herein. For example, in one aspect, the oligonucleotide binds to SEQ ID NO:10 and/or SEQ ID NO:11, wherein the oligonucleotide is between 10 and 50 nucleotides in length. The oligonucleotide can be, for example, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, or 50 nucleotides in length.

[0050] In certain aspects, the oligonucleotide includes at least 10, 11, 12, 13, 14, 15, 16, 17, or 18 contiguous nucleotides of at least one of SEQ ID NOS:1 to 5. In fact, in specific examples of the invention, the oligonucleotide is any one of SEQ ID NOS:1 to 5. Furthermore, the oligonucleotide can include a detectable label.

[0051] In certain embodiments pairs of oligonucleotides disclosed herein form primer pairs, which themselves form another embodiment of the present invention, as disclosed in more detail above.

[0052] Methods of the present invention for determining gender of a canine subject are useful, for example as a quality control step in applications that utilize canine genotyping. As indicated herein, data provided by the methods can be used to verify gender information provided by to a canine registry. Accordingly, the present invention provides a method for quality control testing of a nucleic acid sample from a canine subject, comprising determining gender of the canine subject using a gender marker, and comparing the determined gender with a reported gender of the canine subject. The reported gender can be gender information provided to a registry. In other aspects, gender information determined using methods herein is compared to known gender information. In these embodiments, the

method provides a quality assurance test that an amplification reaction, for example provides valid results.

[0053] Furthermore, aspects of the methods of the invention disclosed herein, wherein human samples generate a different amplification product than canine samples, are useful to confirm that an amplification product(s) was generated from nucleic acids from a canine sample and not from contaminating human DNA. In this aspect, the gender determination is performed as a quality control in a genotyping reaction of at least one microsatellite locus, as disclosed in more detail herein. Accordingly, the quality control testing and the genotyping can be performed in the same reaction using multiplex procedures disclosed herein.

[0054] In another embodiment, provided herein is a method to detect binding of at least one primer or probe to a canine amelogenin gene, that includes contacting a nucleic acid sample from a canine subject with at least one probe or primer specific for canine amelogenin. Embodiments directed at detecting binding of at least one primer or probe to a canine amelogenin gene are identical to those disclosed above for determining gender, except that the methods in certain aspects, do not include the step of determining gender. These methods are useful, for example, for genotyping methods wherein differences among canine subjects in the amelogenin gene are identified and used to identify the canine subject.

[0055] Thus, certain examples of methods for detecting binding of at least one probe or primer to the amelogenin gene are genotyping methods. Accordingly, in certain embodiments, the present invention provides a method to genotype a canine subject, that includes contacting a nucleic acid sample from the canine subject with at least one probe or primer specific for canine amelogenin, and detecting binding of the at least one probe or primer, thereby genotyping the canine subject. These genotyping include probes and primer pairs that specifically bind to nucleotides that are polymorphic and/or probes and primer pairs that provide gender information, as disclosed above.

[0056] For example, in certain aspects the nucleic acid sample is contacted with at least one probe or primer that specifically binds SEQ ID NO:22 and/or SEQ ID NO:23. In another example, the at least one probe or primer specifically binds SEQ ID NO:10 and/or

SEQ ID NO:11. In another example, the nucleic acid sample is contacted with a primer pair, wherein a first primer of the primer pair specifically binds to SEQ ID NO:6 and/or SEQ ID NO:7 and a second primer of the primer pair specifically binds to SEQ ID NO:8 and/or SEQ ID NO:9.

[0057] In certain aspects, for example, the first primer includes at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22 contiguous nucleotides of SEQ ID NO:3, and the second primer includes at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22 contiguous nucleotides of SEQ ID NO:5. In another aspect, the first primer includes at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or 18 contiguous nucleotides of SEQ ID NO:4, and the second primer includes at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22 contiguous nucleotides of SEQ ID NO:5.

[0058] As indicated herein, in certain aspects, methods of the present invention are multiplex reactions. Accordingly, "at least one probe or primer" can include, for example 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or more probes, primers, or combinations thereof. The probes or primers can include probes and primers that specifically hybridize to the amelogenin gene and, in certain aspects, include probes and primers that specifically hybridize to a microsatellite locus (See below). As indicated above, specific hybridization or specific binding is hybridization or binding under stringent hybridization conditions.

[0059] In certain aspects and embodiments disclosed herein, in addition to detecting binding of at least one probe or primer to a canine amelogenin gene, methods of the present invention include detecting binding of at least one probe or primer to a canine microsatellite locus. Canine microsatellite loci and methods for genotyping and identifying additional canine microsatellites are disclosed in U.S. Pat. No. 5,874,217, incorporated by reference in its entirety herein.

[0060] Microsatellites are perfect, imperfect, or compound arrays of tandemly repeated nucleotide sequences embedded in an otherwise unique nucleotide sequence. Microsatellite repeats typically range from one to six base pairs (bp) in length. The microsatellite repeat arrays vary in the number of repeats from 6 to 30 or more in humans. However, the longer

arrays of repeats are less frequently isolated. Microsatellites may consist of simple repeats containing only one uninterrupted repeated sequence, imperfect repeats containing two identical repeats separated by a small interval of non-repeated nucleotides, or compound repeats containing several different repeated sequence types (Weber, 1990). For an individual, any particular microsatellite chromosomal locus may vary in the number of repeats present.

[0061] Commonly used methods of genetic mapping by microsatellites take advantage of length variations among individuals (Weber and May, 1989). Utilizing the nucleotide sequence of the cloned microsatellite and its flanking regions, oligonucleotide primers for the PCR are designed that anneal to unique sequences that flank the repeats. The primers can be designed as near or far from the microsatellite as desired, the only limit being the resulting size of the PCR product for subsequent analysis.

[0062] Methods for detection and size determination of PCR products from a specific microsatellite locus are known in the art, and can be accomplished, for example using amplification methods discussed herein.

[0063] As disclosed in more detail in U.S. Pat. No. 5, 874,217, microsatellite loci that are useful in the subject methods will have the general formula:

$U(R)_nU'$, where

U and U' are non-repetitive flanking sequences that uniquely identify the particular locus, R is a repeat motif, and n is the number of repeats, which is variable in a population. The locus will be present on a canine chromosome. Specific examples are provided in the Sequence Listing, SEQ ID NO: 1 through SEQ ID NO:20.

[0064] Within the flanking sequences U and U', sequences can be selected for amplification primers, P and P'. Tables 4 through 6 provide examples of microsatellite amplification primers. It will be understood that these primers and probes, as well as the canine amelogenin primers and probes disclosed herein are exemplary, and other primers and probes can be identified using the present disclosure and known methods. The number of repeats at a specific locus, n, are polymorphic in a population, thereby generating

individual differences in the length of DNA that lies between U and U'. The number will vary from at least 1 repeat to as many as about 150 repeats.

[0065] Methods that detect binding of probes and primers to both the amelogenin gene and to microsatellite loci are typically multiplex reactions, such as multiplex amplification reactions. Multiplex genotyping provide a genetic fingerprint for the identification of the canine. Multiplex amplification may be performed in which several sets of primers are combined in the same reaction tube. This is particularly advantageous when limited amounts of sample DNA are available for analysis. Conveniently, each of the sets of primers can be labeled with a different fluorochrome or one of a series of fluorochromes included in a reaction.

[0066] For microsatellite analysis, the size of the amplification product is proportional to the number of repeats (n) that are present at the locus specified by the primers. The size will be polymorphic in the population, and is therefore an allelic marker for that locus. For many applications, a panel of loci will be analyzed, usually at least about 5 loci, more usually at least about 10 loci. The combined information for the panel of loci serves as a "DNA fingerprint" for the individual.

[0067] The DNA fingerprint produced by the combination of microsatellite analysis and analysis of the canine amelogenin gene allows for identification of lost or stolen animals and confirmation of identity, such as may be required for show or stud dogs. Genotypes can be compared to assess the degree of relatedness between two individuals. Individuals that are closely related have a higher probability of shared alleles than unrelated individuals (see Chakraborty and Lin (1993) Hum. Biol. 65:875-895; and Hammond (1994) Am. J. Hum. Genet. 55:175-189). Knowledge of relatedness is useful for the improvement of breeding programs, and for increasing genetic diversity in a population. Canine DNA samples recovered from crime scenes may be used to provide forensic evidence.

[0068] The genotypes are useful for parentage testing. Individuals can be ruled out as possible sires or dams of an offspring in question by comparing their genotypes. Each individual must receive one allele from its mother and one from its father at each locus, so that putative parents that do not share an allele at every locus with the individual in question

can be excluded from parentage. When using a set of microsatellite markers for parentage testing, it is important to know how powerful they are in terms of detection of false parentage. The formula for parentage exclusion (PE) states the probability that a given series of codominant alleles should detect a falsely recorded parent, Jamieson (1994) Anim. Genet. 25:37-44. The PE can be calculated for each locus, and for all of the loci combined.

[0069] In embodiments that include analyzing binding of at least one probe or primer to the canine amelogenin gene as well as binding of at least one probe or primer to a microsatellite locus, probes, primers and labels can be selected such that probes, primers, and labels can be used in a single multiplex analysis. For example, as illustrated in the Example herein, primers for the amelogenin gene were chosen that could be used in a 5 dye multiplex system for canine microsatellite markers. Therefore, primers were chosen that generated amplification products (i.e. markers) that did not overlap existing markers (120-156 bases) for a given label (See Figs. 2-4).

[0070] Accordingly, minor changes were made to an initially discovered primer set (BFK90055 and BFK90052) in order to fit them into a canine microsatellite multiplex assay (Figs. 1-4). BFK90065 which is 4 bases smaller than BFK90055, was used in place of BFK90055, and BFK90059 was used in place of BFK90052. BFK90059 is identical in sequence to BFK90052 but fluorescently labeled with NED instead of PET (Applied Biosystems). As reported herein, gender was tested using these primers in over 10,000 dog samples, and in all of these assays, gender was correctly identified. Examples of gender determination by the analysis of nucleic acid products derived from the use of canine amelogenin specific primers are illustrated in Figures 1-4.

[0071] Aspects of the present invention which include gender determination using the amelogenin gender markers disclosed herein, as well as genotyping using canine microsatellite markers, provide for the both a determination of gender and identity of a canine subject. This has the advantage of efficiently providing this information using an automated multiplexing system, of which many are known in the art.

[0072] Accordingly, provided herein is a method for simultaneously determining gender, identity and parentage in a canine, that includes contacting a nucleic acid sample from a

canine to be tested, with primers specific for canine amelogenin as disclosed herein, and probes and primers specific for DNA-based parentage and identity (e.g. microsatellite locus) in a multiplex reaction, and identifying the nucleic acid products, thereby determining gender, identity and parentage. This method is illustrated in the Example and Figures 2-4 provided herein. In other words, one embodiment the present invention provides a method and system, also referred to as a "comprehensive canine testing system," whereby the gender marker is included in a multiplex reaction that includes other markers, e.g., microsatellite markers, that provide the basis for identity and parentage testing. Therefore, the method can be multiplex genotyping reactions that utilize the gender markers disclosed herein, as well as short tandem repeats (STRs), also referred to as microsatellites. Microsatellite markers offer numerous benefits, including that these markers are evenly distributed across the genome, are highly polymorphic, and therefore, highly informative.

[0073] Microsatellite genotyping and gender data regarding a canine subject can be stored in a computer readable form, such as a database. Accordingly, in certain aspects of the invention, a dog owner provides a sample to a service provider for genotyping and gender analysis. For example, the sample can be sent by mailing the sample to the service provider. The service provider can then perform a gender determination and microsatellite genotyping, for example in a single multiplex reaction, as disclosed herein, and store the results in a database.

[0074] The database can be flat file databases, relational databases or object-oriented databases. The parts of the database can be internal databases, or external databases that are accessible to users. An internal database is a database maintained as a private database, typically maintained behind a firewall, by an enterprise. An external database is located outside an internal database, and is typically maintained by a different entity than an internal database.

[0075] Amelogenin gender markers are markers that are detectably different when generated from the X chromosome compared to the Y chromosome. For example, amelogenin gender markers include amplification products generated using primers disclosed herein that produce a different sized product when used to amplify a portion of the

amelogenin gene on the X chromosome as compared to the product generated from the amelogenin gene on the Y chromosome. Amelogenin gender markers can also include restriction fragments that are detected using a probe of the present invention, that are a different size in the X chromosome amelogenin gene than the Y chromosome amelogenin gene. Microsatellite markers are markers that are detectably different depending on the number of repeats at a microsatellite locus.

[0076] In certain aspects, the microsatellite locus includes a repeated tetranucleotide motif consisting of AAA or TTT, with the fourth residue of the tetranucleotide repeat motif being any one of G, C, A or T. In another aspect, the microsatellite locus comprises a repeated motif selected from AAAG, GAAA, AAAT, TTTC, CTTT, TTTA, AAGG, GAAT, GAAG, GAAAA, AAAAAG, TGC and TTC.

[0077] In certain aspects the microsatellite locus is at least one of PEZ1/CATA1, PEZ3, PEZ5, PEZ6, PEZ8, PEZ10, PEZ11, PEZ12, PEZ13, PEZ15, PEZ16, PEZ17, PEZ20, PEZ21, FH2010 ("FH" denotes a Fred Hutchinson marker), FH2054, and FH2079 (See Tables 4-6). Accordingly, for example, binding of at least one primer or probe to each of PEZ1/CATA1, PEZ3, PEZ5, PEZ6, PEZ8, PEZ12, PEZ20, FH2010 (Fred Hutchinson markers), FH2054, and FH2079, or any combination thereof, is detected. For example, a panel of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 of the microsatellites can be genotyped. In certain aspects, this method is a multiplex amplification reaction that employs a primer pair for generating a different sized amplification product for the X chromosome amelogenin gene from that obtained for the Y chromosome amelogenin gene, as disclosed herein, as well as the primer pairs listed in Tables 4 through 6 for each of PEZ1/CATA1, PEZ3, PEZ5, PEZ6, PEZ8, PEZ12, PEZ20, FH2010, FH2054, and FH2079. Thus, this embodiment is a 10-plex satellite marker panel along with an amelogenin gender marker.

[0078] In another example, binding of at least one primer or probe to each of PEZ10, PEZ11, PEZ13, PEZ15, PEZ16, PEZ17, and PEZ21, or any combination thereof, is detected. For example, this method can be a multiplex amplification reaction that employs a primer pair for generating a different sized amplification product for the X chromosome

amelogenin gene from that obtained for the Y chromosome amelogenin gene, as disclosed herein, as well as the primer pairs listed in Tables 4 through 6 for each of PEZ10, PEZ11, PEZ13, PEZ15, PEZ16, PEZ17, and PEZ21. Thus, this embodiment is a 7-plex satellite marker panel along with an amelogenin gender marker.

[0079] In yet another example, binding is detected for at least one primer or probe to each of PEZ1/CATA1, PEZ3, PEZ5, PEZ6, PEZ8, PEZ12, PEZ16, PEZ17, PEZ20, PEZ21, FH2010, FH2054, and FH2079, or any combination thereof. For example, this method can be a multiplex amplification reaction that employs a primer pair for generating a different size amplification product for the X chromosome amelogenin gene from that obtained for the Y chromosome amelogenin gene, as disclosed herein, as well as the primer pairs listed in Tables 4 through 6 for each of PEZ1/CATA1, PEZ3, PEZ5, PEZ6, PEZ8, PEZ12, PEZ16, PEZ17, PEZ20, PEZ21, FH2010, FH2054, and FH2079. Thus, this embodiment is a 13-plex satellite marker panel along with an amelogenin gender marker.

[0080] In another embodiment, the present invention provides a kit for performing the methods provided herein. Such a kit can include, for example, at least one set of primers useful for amplifying canine amelogenin and optionally at least one other pair of primers useful for amplifying microsatellite DNA repeats. In certain aspects the kit includes a pair of primers useful for amplifying canine amelogenin and a panel of primers useful for amplifying a panel of microsatellite DNA repeats. The primers can be conjugated to a detectable marker, and can be conjugated to different labels to facilitate multiplex amplification. In a preferred embodiment, the primers include SEQ ID NOS: 3 or 4 and SEQ ID NO:5.

[0081] The present invention discloses the nucleotide and polypeptide sequence of a portion of the canine amelogenin gene, which as indicated above, can be used to identify the nucleotide sequence of the entire canine (e.g. dog) amelogenin X chromosome gene and Y chromosome gene, as well as the polypeptide sequence of canine amelogenin from the X chromosome and the Y chromosome. Accordingly, in another embodiment, the present invention provides an isolated polynucleotide that includes all or a portion of a canine (e.g.

dog) amelogenin gene, for example at least 25, 50, 75, 100, 150, 200, 250, 300, 400, 500, 750, or 1000 contiguous nucleotides of a canine amelogenin gene, or a complement thereof, or a polynucleotide at least 75, 80, 85, 90, 95, 96, 97, 98, 99, or 99.9% identical thereto. In one aspect, the present invention provides an isolated polynucleotide that includes at least 25, 50, 75, 100, 150, 200, 250, 300, 400, or 492 contiguous nucleotides of SEQ ID NOS:22 or 23, or a complement thereof, or an isolated polynucleotide that is 85, 90, 95, 96, 97, 98, 99, or 99.9% identical to these sequences; or an isolated polynucleotide that includes at least 400, 450, 500, 550, 600, 650, or 680 contiguous nucleotides of SEQ ID NO:23, or a complement thereof, or an isolated polynucleotide that is 85, 90, 95, 96, 97, 98, 99, or 99.9% identical to these sequences. In certain aspects, the present invention provides an isolated polypeptide that includes SEQ ID NO:22 or SEQ ID NO:23. SEQ ID NOS:22 and 23 provide the nucleotide sequence of a portion of the canine amelogenin gene on the X and Y chromosomes, respectively. It will be understood that based on the present disclosure, routine methods can be used to identify the entire canine amelogenin gene.

[0082] The amelogenin gene is involved in the production of tooth enamel. Therefore, in addition to uses of the amelogenin gene and polypeptide (see below) sequences in gender determination, these compositions are also useful for identifying new methods to diagnose and treat canine tooth disorders. Furthermore, these compositions can be used for improved tooth enamel formation.

[0083] In another embodiment, the present invention provides a vector comprising the isolated polynucleotide portion of the canine amelogenin gene disclosed above. The term "vector" refers to a plasmid, virus or other vehicle known in the art that has been manipulated by insertion or incorporation of a nucleic acid sequence.

[0084] Methods that are well known in the art can be used to construct vectors of the present invention, including *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* recombination/genetic techniques (See, for example, the techniques described in Maniatis et al. 1989 Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y., incorporated herein in its entirety by reference).

[0085] In another aspect, the present invention provides an isolated cell that includes a vector of the present invention. The cell can be prokaryotic or eukaryotic. Techniques for incorporated vectors into prokaryotic and eukaryotic cells are well known in the art. In certain aspects, the cells are canine cells. In other aspects, the cells are bacterial cells. In still other aspects, the cells are human cells.

[0086] In another embodiment, the present invention provides an isolated polypeptide encoded by SEQ ID NO:22 or SEQ or SEQ ID NO:23. Furthermore, in certain embodiment, the present invention provides an antibody that specifically binds to the isolated polypeptides of the present invention.

[0087] In another embodiment, the present invention provide an isolated polynucleotide that encodes the polypeptide encoded by SEQ ID NOS:22 or 23, or a complement thereof.

[0088] The following example is intended to illustrate but not limit the invention.

EXAMPLE 1

[0089] DETERMINATION OF CANINE GENDER

[0090] Canine amelogenin primer design

[0091] All canine DNA samples analyzed in this example, were dog DNA samples. Several sets of primers were designed based on similarities among canine, human and other mammalian sequences corresponding to several sites along the amelogenin genes. One set of primers, BFK90004 (SEQ ID NO:2) and BFK90006 (SEQ ID NO:1), when used to PCR amplify male or female canine DNA, produced many bands on an agarose gel although the expected size (somewhat larger than 2000 base pairs, based on human sequence) was very faint from female canine DNA and strong from male canine DNA. Primers BFK90004 (SEQ ID NO:2) and BFK90006 (SEQ ID NO:1) correspond to sequences in the human amelogenin exons 3 and 2 and span intron 2 which is about 1940 bp on the X chromosome and 2500 bp on the Y chromosome. BFK90006 (SEQ ID NO:1), which includes the ATG start codon for the gene, with one degenerate base, is identical within human, equine, bovine and goat X and Y chromosomes and with other mammals where amelogenin chromosomal locations are either unknown or only X-linked, such as porcine and mouse.

BFK90004 (SEQ ID NO:2) is the reverse complement of one of the primers of a set that was used to amplify canine fragments successfully and is identical in human X and Y chromosomes.

[0092] Amplification and cloning of canine amelogenin DNA

[0093] DNA was isolated from various male and female dogs using either the PUREGENE DNA Isolation Kit for blood (Gentra Systems) or a chelex prep for buccal swabs (Walsh et al. 1991). PCR was performed on 300ng of male or female canine genomic DNA using 10 pmol each of primers BFK90004 (SEQ ID NO:2) and BFK90006 (SEQ ID NO:1), 1X Taq Gold Buffer (Applied Biosystems), 2mM MgCl₂, 0.2mM each dNTP (Applied Biosystems) and 2 units of AmpliTaq Gold Polymerase (Applied Biosystems) in a 50 µl reaction. The following PCR conditions were used in a Applied Biosystems 9700 thermocycler: 95°C for 10 minutes followed by 35 cycles of 30 seconds at 94°C, 30 seconds at either 56°C or 58°C, and 1 minute 30 seconds at 72°C followed by a final extension of 72°C for 20 minutes. 25 µl of each product was fractionated alongside a low mass DNA ladder (BRL (now Life Technologies)) on a 0.85% SeaPlaque GTG (BioWhittaker) agarose gel in TAE buffer containing GelStar (BioWhittaker) stain. Bands of three different sizes (>2 kb, >1.2 kb and <1.2 kb) were excised from the gel, purified with a QIAquick Gel Extraction Kit (Qiagen) and eluted in 30 µl. 2 µl of each DNA fragment was ligated into a pCR'4-TOPO vector using the TOPO TA Cloning Kit for Sequencing (Invitrogen), electroporated into TOP 10 cells (Invitrogen) and plated on agar plates containing 50 µg/ml kanamycin. For each construct, 16 resulting bacterial colonies were individually picked by touching with a sterile toothpick and lightly swabbing into the bottom of a 200 µl PCR tube. The inserts were amplified in a 20 µl reaction containing 5 pmol each of M13 forward and reverse primers, 1X Taq Gold Buffer, 2mM MgCl₂, 0.2mM each dNTP and 1 units of AmpliTaq Gold Polymerase in an Applied Biosystems 9700 thermocycler using the conditions above for 30 cycles instead of 35 and with a 56°C annealing temperature. To test for inserts and sizes, 5 µl of the reactions were fractionated on a 0.8 % agarose gel.

[0094] Sequencing

[0095] For clones containing inserts, 2.5 µl of the PCR product was removed to a tube containing 2 µl of a solution containing 1 unit each of Exonuclease I and Shrimp Alkaline Phosphatase (USB) per microliter in water and incubated at 37°C for one hour followed by inactivation at 75°C for 20 minutes. Sequencing was carried out with a BigDye Terminator V3.0 Cycle Sequencing Kit (Applied Biosystems) using 2 µl of the treated fragments and 3.2 pmol of either T3 or T7 primer from the TA cloning kit (Invitrogen). Reactions were ethanol precipitated, resuspended in 20 µl of Hi-Di Formamide, denatured and run on an ABI Prism 3100 using POP6 and a 50 inch capillary array. Resulting sequences were analyzed and compared using Lasergene software (DNASTAR).

[0096] The constructs of the expected size (>2 kb) produced sequences consistent with an amelogenin coding region. Based on differences within our canine sequences, new primers were developed.

[0097] Gender marker primer design and fragment detection

[0098] Sequences that were identical between the putative X and Y canine genes and that spanned insertion/deletion areas were chosen for primer design with the aim of producing fragments of different sizes for each chromosome. One region of the canine amelogenin gene that was identified as useful in gender determination is shown in Tables 1 and 3.

[0099] Primers were chosen such that they could be used in a 5 dye multiplex system for canine microsatellite markers. Therefore, primers were chosen that generated amplification products (i.e. markers) without overlapping existing markers (120-156 bases). Furthermore, primer sets were chosen would not amplify human DNA, which would be a potential source of contamination. One primer of each set was fluorescently labeled. Exemplary primers that were identified are shown in Table 2.

[0100] Male or female canine DNA (10 ng) was each amplified in an 8 µl reaction using 0.5 pmoles of each primer, 1.4X Taq Gold Buffer, 3 mM MgCl₂, 0.27 mM each dNTP and 0.18 units of AmpliTaq Gold Polymerase in a Perkin Elmer 9600 thermocycler. PCR conditions were: 95°C for 12 minutes followed by 35 cycles of 30 seconds at 94°C, 30

seconds at 58°C, and 1 minute at 72°C followed by a final extension of 72°C for 30 minutes. Following PCR the samples were diluted 30 fold and 2 µl were added to 10 µl Hi-Di Formamide containing 0.1µl GeneScan -500 LIZ Size Standard (Applied Biosystems), denatured and run on an ABI Prism 3100 using POP4 and a 36 inch capillary array and analysed using Genotyper software (Applied Biosystems). One set of primers (BFK90052 and BFK90055) produced a 143 base band in male and female canine samples and an extra band of 146 bases in the male sample. These results were consistent across 3 male and 3 female dog samples.

[0101] Minor changes were made in the primer set in order to fit them into a canine microsatellite multiplex assay (Figures 2-4). BFK90065 which is 4 bases smaller than BFK90055, was used in place of BFK90055, and BFK90059 was used in place of BFK90052. BFK90059 is identical in sequence to BFK90052 but fluorescently labeled with NED instead of PET (Applied Biosystems). Examples of gender determinations using these primers are provided in Figure 1. Two amplification products (denoted "X" and "Y" in Figures 1-4), were generated using male canine subjects. Amplification products generated from female subjects, on the other hand, included only one product (denoted "X" in Figures 1-4). Gender has been tested using these primers in over 10,000 dog samples from a wide range of breeds, and in these assays, gender was correctly identified.

[0102] Assay of Multiplexed PCR Primers for Detection of Polymorphism:

[0103] PCR was performed on canine genomic DNA using microsatellite markers pooled with the gender marker using the GeneAmp™ PCR process in a 9600 thermocycler (Perkin Elmer, Norwalk CT). Post-PCR products were run on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, Foster City CA). GS500LIZ (Applied Biosystems, Foster City CA) markers were present in each sample as internal size standards. PCR products were sized and genotyped using GeneScan™ Analysis Software and Genotyper™ Software (Applied Biosystems, Foster City CA). Primers for microsatellite genotyping are listed in Tables 4 through 6. More specific PCR conditions are listed below. Examples of genotyping results are illustrated in Figures 2 through 4. Allele products are labeled with the marker name and size in base pairs. Alleles for the gender marker are labeled as X or Y

with base pair sizing. Figure 2 is from a female dog and Figures 3 and 4 are from male dogs. The results illustrated in Figures 2-4 establish that methods of the present invention can provide both genotype and gender information in a single reaction.

[0104] Specific example of a method of the invention

[0105] A specific set of reaction and PCR conditions that can be used for gender marker identification using methods for determining gender disclosed herein, within a multiplex genotyping reaction disclosed above, are as follows: 0.29 pmoles of each of BFK90059 and BFK90065 primers, 1.5X Taq Gold Buffer, 2.5 mM MgCl₂, 0.175 mM each dNTP and 0.75 units of AmpliTaq Gold Polymerase in an 5 µl reaction containing a highly variable amount of canine DNA (approximately 0.5-100 ng). PCR: 95°C for 12 minutes, 15 cycles of 30 seconds at 94°C, 30 seconds at 64°C decreasing 0.2°C/cycle to 61°C, and 1 minute at 72°C, followed by 20 cycles of 30 seconds at 94°C, 30 seconds at 61°C, and 1 minute at 72°C, followed by a final extension of 72°C for 45 minutes in a Perkin Elmer 9600 thermocycler.

[0106] Sequence Support

[0107] The sequence data provides evidence that the disclosed primers are specific for the canine amelogenin gene. Table 7 provides an alignment of sequence information obtained from canine compared to human amelogenin sequences. When human DNA was tested using primers BFK90065 and BFK90059, a faint 143 base band was detected in female and male but no 146 base band was detected in any human sample thus eliminating any concern of human DNA contamination. These results were expected since BFK90065 is in a highly conserved region near the initial amelogenin ATG start codon whereas BFK90059 lies within an intron and has 19/22 bases identical to human amelogenin-X but only 15/22 bases identical to human amelogenin-Y with the 3' end of the primer being most different (Table 1).

[0108] Table 1: Comparison of Canine gender marker sequence with Human amelogenin sequence (italics denote primer binding region and bold denotes differences from consensus) (contiguous human sequences from table below are found in SEQ ID NO:20 (X

chromosome) and SEQ ID NO:21 (Y chromosome) and contiguous canine sequences are found in SEQ ID NO:10 (X chromosome) and SEQ ID NO:11 (Y chromosome).

	SEQ ID NO:	
Human X	12	GATTTTATTTGCCTGCCTCCTGGGAGCAGCTTTTGCCATGCCTGTGAGTAAAACACCCCTTTCATAAGTCAGTGTC
Human Y	13	GATTTTGTTTGCCTGCCTTGTGGGAGCAGCTTTTGCCATGCCTGTGAGTAAAACATTCTTTGCATAAGTTG-TGTC
Canine X	14	GATTTTGTTTGCCTGCCTCCTGGGAGCAGCCTTTTGCTATGCCCTGAGTAAAATACCCCTG-CATAA-T-A-T-T-
Canine Y	15	GATTTTGTTTGCCTGCCTCCTGGGAGCAGCCTTTCAGTATGCCTGTGAGTAAAAT-----TT-CC-AA-T-T-T-C- BFK90065 →
Human X	16	CAATTCACAACTTGGACATAAAAAATCTGC-TCATA-----GTT-GG-TGAAA-TT-AGGGTTTAAAACAGTATGAGATC
Human Y	17	CAATTT-ACAAATGGAATAAGAATTTG-TC--CCCC-AGCT-GG-GAACTTTCC-AAGGTTTAAAACAGTTTGAGCTA
Canine X	18	CAATTCACAAGCTTGGAATAAAAAGTCTGCC-CC-AC----AGTT-GG-T-AAA-CTTTAGGGTTTAAAGACAGTACAAGATC
Canine Y	19	CAATTCACAAGCTTGGAATAAAAATCTGCCTCAGATTTTATTATTAGGGTTAAA-TTTAAGGGTTTAAAGACAGTACAAGATC ← BFK90059

[0109] Table 2. Canine amelogenin gender marker primers and sequence data

Primer	Sequence	Comments	SEQ ID NO:
BFK90006	5'ATGGGGACCTGGATTTRTTTG 3'	(R= A or G)	1
BFK90004	5' CCAGGGTGCCCAGGATGAGGT 3'		2
BFK90055	5' GATTTTGTTTGCCTGCCTCCTG 3'		3
BFK90065	5' GATTTTGTTTGCCTGCCT 3'	(BF0055 - 4 bases)	4
BFK90052	5' GATCTTGTACTGTCTTAAACCC 3'	(DYE-LABELED A)	5
BFK90059	5' GATCTTGTACTGTCTTAAACCC 3'	(DYE-LABELED B)	5

[0110] Table 3. Canine Gender Marker Sequences (*italics* denotes primer binding region and **bold** denotes differences in sequence)

	SEQ ID NO:	
Canine X (1-76)	6	<i>GATTTTGTTCCTGCCTCCTGGGAGCAGCCTTTGCTATGCCCGTGAGTAAATACCCCTGCATAATTTCAATTT</i>
Canine Y (1-70)	7	<i>GATTTTGTTCCTGCCTCCTGGGAGCAGCCTTCAGTATGCCTGTGAGTAAAT-----TTCC-AATTTCCAATTT</i> BFK90065 →
Canine X (77-143)	8	CACAAGCTTGGAATAAAAGTCTGCC-CC-ACC-----AGTT-GG-T- AACTTTAGGGTTTAAGACAGTACAAGATC
Canine Y (71-146)	9	CACAAGCTTGGAATAAAATCTGCCTCAGATTTTATTAGGGTTAAATTAAAGGGTTAAGACAGTACAAGATC ← BFK90059

[0111] Table 4. Microsatellite primers and sequences

Primer 1 SEQ ID NO:	Primer 2 SEQ ID NO	Locus ID SEQ ID NO	PEZ Nomenclature	
52	72	32	PEZ3	
53	73	33	PEZ6	
54	74	34	PEZ7	
55	75	35	PEZ8	
56	76	36	PEZ9	
57	77	37	PEZ5	
58	78	38	PEZ10	
59	79	39	PEZ11	
60	80	40	PEZ13	
61	81	41	PEZ20	
62	82	42	PEZ14	
63	83	43	PEZ21	
64	84	44	PEZ15	
65	85	45	PEZ16	
66	86	46	PEZ19	
67	87	47	PEZ22	
68	88	48	PEZ12	
69	89	49	PEZ18	
70	90	50	PEZ17	
71	91	51	PEZ2	

[0112] Table 5. FHC Microsatellite Markers (all are tetra repeat motifs)

Locus	Size range	Chromos. Location	Forward	SEQ ID NO:	Reverse	SEQ ID NO:
FH2054	135-179	CFA12	GCCTTATTCATTG CAGTTAGGG	92	ATGCTGAGTTTTGAACCTT TCCC	95
FH2079	263-299	CFA24	CAGCCGAGCACA TGGTTT	93	ATTGATTCTGATATGCCC AGC	96
FH2010	220-248	CFA24	AAATGGAACAGT TGAGCATGC	94	CCCCTTACAGCTTCATTT TCC	97

[0113] Table 6. Microsatellite Marker PEZ1/CATA1 (tetra repeat motif)

Locus	Size range	Chromos. Location	Forward	SEQ ID NO:	Reverse	SEQ ID NO:
PEZ1/ CATA1	95-133	CFA07	GGCTGTCACCTTTT CCCTTTC	98	CACCACAATCTCTCTCAT AAATAC	99

[0114] Table 7. Comparison of Human and Canine X and Y amelogenin sequences and below them, the predicted proteins. Bolded amino acids indicate difference from consensus. The italicized sequences indicates the beginning of an intron known in human and assumed in canine.

	SEQ ID NO:	
Human X	24	ATGGGGACCTGGATTTTATTTGCCTGCCTCCTGGGAGCAGCTTTTGCCATGCCTGTGAGTAAAA
	25	M G T W I L F A C L L G A A F A M P
Human Y	26	ATGGGGACCTGGATTTTGTTCCTGCCTGTGGGAGCAGCTTTTGCCATGCCTGTGAGTAAAA
	27	M G T W I L F A C L V G A A F A M P
Canine X	28	ATGGGGACCTGGATTTTGTTCCTGCCTCCTGGGAGCAGCCTTTGCTATGCCCGTGAGTAAAA
	29	M G T W I L F A C L L G A A F A M P
Canine Y	30	ATGGGGACCTGGATTTTGTTCCTGCCTCCTGGGAGCAGCCTTCAGTATGCCTGTGAGTAAAA
	31	M G T W I L F A C L L G A A F S M P

[0115] REFERENCES

[0116] Buel E, Wang G, Schwartz M. (1995) PCR Amplification of Animal DNA with Human X-Y Amelogenin Primers Used in Gender Determination. J Forensic Sci 40:641-644.

[0117] Chen CM, Hu CL, Wang CH, Hung CM, Wu HK, Choo KB, Cheng WT. (1999) Gender determination in single bovine blastomeres by Polymerase chain reaction

amplification of sex-specific polymorphic fragments in the amelogenin gene. *Mol Reprod Dev* 54(3):209-214.

[0118] Hasegawa T, Sato F, Ishida N, Fukushima Y, Mukoyama H. (2000) Sex Determination by Simultaneous Amplification of Equine SRY and Amelogenin Genes. *J Vet Med Sci* 62(10):1109-1110.

[0119] Sullivan KM, Mannucci A, Kimpton CP, Gill P. (1993) A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. *Biotechniques* 15:636-641.

[0120] Walsh PS, Metzger DA, Higuchi R. (1991) Chelex 100 as a medium for simple extraction of DNA for PRC-based typing from forensic material. *Biotechniques* 10:506-518.

[0121] Griffiths, R., Daan, S., Dijkstra, C. (1996) Sex identification in birds using two CHD genes. *Proc. R. Soc. Lond. B Biol. Sci.* 263(1374):1251-6.

[0122] Although the invention has been described with reference to the above example, it will be understood that modifications and variations are encompassed within the spirit and scope of the invention. Accordingly, the invention is limited only by the following claims.